DOCUMENT RESUME

ED 262 069 TM 850 546

TITLE	An Update on the Alaska Statewide Testing Program. Assessment Reports 8-10.
INSTITUTION	Alaska State Dept. of Education, Juneau, Office of Education, Assessment and Research.
PUB DATE	85
NOTE	30p.; For Assessment Report 1-7, see ED 242 736.
AVAILABLE FROM	Alaska Assessment Materials, Interwest Applied Research, P.O. Box 8189, Portland, OR 97207.
PUB TYPE	Collected Works - Serials (022) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Elementary Education; *Evaluation Methods; Grade 4; Grade 8; *Mathematics Skills; *Reading Skills; *Standardized Tests; *State Programs; Test Construction; *Testing Programs; Test Interpretation; Test Results; Test Wiseness
IDENTIFIERS	*Alaska Statewide Assessment Test Program

ABSTRACT

The Alaska Statewide Assessment Test (ASAT) in reading and math is administered to fourth and eighth graders every two years. This document contains three separate assessment reports generally titled "An Update on the Alaska Statewide Testing Program." Assessment Report 8, "The Interpretive Panels Look at Results," lists student strengths and weaknesses on the 1983 ASAT based on a review of the test content by 70 educators from throughout the state who were nominated by their districts to serve as interpretive panel members. Assessment Report 9, "What Alaska Students Should Learn: Perspectives from Those Who Teach Them," compares results from the fourth statewide assessment (1983) with curricular emphases recommended by the Alaska Council of Teachers of Mathematics and the Alaska State Reading Association. Assessment Report 10, "Better Assessment Scores are Possible. A Test Author Tells How...," discusses the preparation for the March 1985 ASAT, steps to improve students' scores, previously assessed reading and math weaknesses (1983), and questions and answers regarding assessment policy. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED262069

Assessment Reports 8-10:

An Update on the Alaska

Statewide Testing Program

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Hazelton, Alexander

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 850 544

ASSESSMENT REPORT 8: An Update on the ALASKA STATEWIDE TESTING PROGRAM

The Interpretive Panels Look at Results

In March 1983, more than 13,000 Alaska fourth and eighth graders participated in the state's fourth statewide assessment. This assessment, administered once every two years, measures student achievement in basic skill areas considered significant by Alaska educators and content specialists.

The results of the March test have been available for several months; they were reported in *Assessment Report 7*. However, as indicated previously, there was no way to judge student performance—as satisfactory, or above or below expectations—until the results were compared to the expectations held for the students by Alaska's educators.

The reviews and analyses needed to make such comparisons have been completed; these findings form the heart of this report.

HOW THE INTERPRETIVE PANELS WORK

Last year, district superintendents and members of the Content Panels who assist with test development and review nominated approximately 70 educators—each familiar with reading or mathematics curricula—to serve as Interpretive Panel members.

These panel members each reviewed one of four test sections—grade 4 math, grade 8 math, grade 4 reading, or grade 8 reading—in order to provide their expectations of student performance on each item. In making their judgments, panelists carefully considered

- The difficulty of each item, relative to others which measured the same skill,
- The importance of the skill being measured (as both a skill in itself and as a foundation for higher skills), and
- The expectations they hold for students at each grade level.

Then, before seeing any actual student results, Interpretive Panel members, working independently, established two levels of expectation:

- First, a MINIMUN level: that is, the lowest percentage of students answering correctly that could be regarded as acceptable.
- Second, a DESIRED level: that is, the highest percentage of students that panel members would realistically expect to answer the item correctly.

Panelists' individual scores were then averaged to produce a final figure against which actual student performance would be judged. When actual performance fell below the Minimum, it was said to indicate a WEAKNESS. When performance exceeded the Desired level, it was said to indicate a STRENGTH. Performance within the range of Minimum to Desired was regarded as SATISFACTORY. Remember that this range of satisfactory performance differs considerably from item to item. Thus, a performance score of, say, 70% might indicate a strength on one item—but a weakness on another.

Because Interpretive Panel members did not see the test results prior to setting the levels, and because they worked independently—not knowing how other Panel members were responding—Assessment Program staff feel confident that the process yields an unbiased indication of the expectations held by Alaska educators.

HOW TO EVALUATE STRENGTHS AND WEAKNESSES

Once the job of the Interpretive Panel is completed, it is up to individual districts to determine how best to use assessment results. It is not within the scope of this report to offer recommendations on using results or to suggest what impact, if any, these results should have on state or district curriculum. But we can present some considerations to assist you in deciding what the results mean.

TM 850 546 (1 of 3)

First, keep in mind that the Alaska Assessment tests are NOT minimum competency tests. The items contained on the assessment tests are of variable difficulty, and it would not be reasonable to expect all students to answer a given item correctly.

Second, because Alaska has no statewide curriculum, each of the state's 53 districts independently determines what content will be taught at a given grade level. Such content may not always match what is covered by the assessment test. Or, even where a match exists, certain skills may not be covered prior to the March administration of the assessment test. Interpretive Panel members keep these issues in mind in setting their levels of expectation. But even so, apparent strengths or weaknesses should be carefully reviewed by those local educators who are most familiar with district curriculum content and goals.

Also remember that items measuring a given skill can vary in difficulty. For example, approximately 60% of the eighth graders tested could identify the main point in a short article about hurricanes. But suppose that item had been made much more difficult? Or much easier? Student performance might have looked considerably different. The items were carefully constructed and reviewed to make certain that they were good indicators of whether students had a specific skill. Nevertheless, you couldn't say that 60% of Alaska eighth graders know how to identify the main point of any written article. Only after the Interpretive Panel looks at the item and sets its expectations can the 60% figure be classed as a "good" or "bad" performance.

While Content Panel members who helped develop and review this test selected items of a content and difficulty suitable for *statewide* testing purposes, district educators should still use their personal judgment when evaluating results. Review copies of the test can be obtained for this purpose at no cost from the Alaska Department of Education. Phone 465-2966.

ONE FINAL POINT . . .

While it is correct to make comparisons between actual student performance and the Interpretive Panels' expectations, it would be highly inappropriate to make any comparisons between fourth and eighth graders' performance. This is because the items for the two grade levels are not equal in difficulty. Even though fourth graders obtained a higher percent correct than eighth graders on the Reading Comprehension subtest, for example, it would not be accurate to say that they performed "better" because the two tests are not comparable. The skills measured are somewhat different, and the relative difficulty for grade level—even within the range of appropriateness defined by the Content Panel—varies substantially.

Would you like to know more about how the Alaska Assessment tests originated, how items were developed and reviewed, or how the tests were administered and scored? These issues and many others are covered in earlier Assessment Reports. Copies are available from the Office of Research and Evaluation, Alaska Department of Education, Pouch F, Juneau 99811.

Remember that the underlying purpose of assessment is to identify instructional needs to assist districts and the state in planning how to improve the quality of Alaska education. This purpose is best achieved through careful comparison of actual performance to the expectations of Alaska educators and the general public.

How did students do on these subtests overall?

The Alaska Assessment tests actually comprise six subtests—three for Mathematics and three for Reading. Math subtests include

- Application
- Computation
- Concepts

Reading subtests include

- Comprehension
- Study Skills
- Word Skills

Fourth graders gave their best overall performance in Reading Comprehension, with five strengths, nine satisfactory performances, and only one weakness. They had the most trouble with the other two Reading subtests—Study Skills and Word Skills. In both cases, identified weaknesses outnumbered strengths and satisfactory performances combined.

For eighth graders, the biggest trouble spots occurred in Math. There were no identified strengths in either Computation or Concepts—and only one in Application. And there were a total of 14 weaknesses identified on the three Math subtests. On the other hand, eighth graders had a total of 10 strengths in Reading—and only seven weaknesses out of a total 42 Reading items.

Now let's consider—by subtest and by grade level—how the Interpretive Panel judged student performance for 1983 on each item.

MATHEMATICS APPLICATION

Application items test problem solving skills: identifying facts, translating written information into math equations, choosing the correct approach to solve a problem, determining which data are needed to solve a problem, and so on. Items generally consist of a short story problem, with one or two relevant questions.

GRADE 4

Summary of Performance
3 STRENGTHS 3 WEAKNESSES

Item Performance

Item 2 Understanding the Problem

Given a word problem requiring addition, 87.2% could answer the question "What does this problem ask you to find?" Range was 66.3 to 84. *A **strength**.*

Remember that the "range" indicated in each case refers to the span between the MINIMUM level and DESIRED level. In Item 2, preceding, the MINIMUM level was 66.3%, while DESIRED was 84%. The 87.2% performance level was above this range, indicating a strength.

Item 3 How to Solve the Problem

In Part 2 of the same problem, 91.8% knew how to use addition to find the total number of marbles owned by three children. Range was 69.7 to 85.1. *A **strength**.*

Item 7 Understanding the Problem

In Part 1 of another two-part problem, 73.6% recognized that they were asked to find the area of a shed, rather than the height, width, or perimeter. Range was 46.9 to 71. *A **strength**.*

Item 8 How to Solve the Problem

But in the second part of the shed problem, only 47.6% knew which formula (4 x 3) would give them the area. Range was 48.9 to 74.3. *A **weakness**.*

Item 9 How to Solve the Problem

Items 9 and 10 were based on a short word problem about Anchorage. Fewer than half the students—44.4%—knew how to use subtraction to determine the year in which Anchorage had become a city. Range was 48.2 to 74.3 *A **weakness**.*

Item 10 Identifying Facts

Item 10 proved difficult as well. Only 36.7% could give the population of Anchorage when it became a city, although this figure was directly stated in the problem. Range was 50.1 to 74. *A **weakness**.*

Comments

Notice that students did best on the two-part problem involving addition (Items 2 and 3). In the problems on area (Items 7 and 8), students knew what was asked for, but many did not know it was necessary to multiply length times width to find area. Problems with Items 9 and 10 may have been caused by hasty reading, since a number of students who responded incorrectly on these items seem to have confused the figures for year and population.

GRADE 8

Summary of Performance
1 STRENGTH 5 WEAKNESSES

Item Performance

Item 2 Averages

This item required students to find the average of three dollar figures involving decimals. 58.9% answered correctly; range was 65.7 to 90. *A **weakness**.*

Item 3 How to Solve the Problem

In the single best math performance by either grade on any item, 97.4% correctly identified adding as the way to solve a short word problem on the number of people visiting the Mt. Hood Glacier. Range was 88.5 to 96.9—relatively high, but performance was still better. *A **strength**.*

Item 12 How to Solve the Problem

53.6% knew how to solve this two-step word problem using subtraction. Range was 57 to 84. *A **weakness**.*

Item 14 Missing Data

63.2% knew which missing data were needed to solve a word problem on averages. Satisfactory range was just a shade higher: 63.6 to 87. *A **weakness**.*

Item 15 Unneeded Data

In this turnabout version of Item 14, only 49.4% could tell which data were NOT needed to solve a word problem requiring multiplication. Range was 58.6 to 82.7. *A **weakness**.*

Item 16 How to Solve the Problem

In Part 2 of the Item 15 word problem, 54.6% could select the correct equation to solve the problem. Range was 57 to 83.6. *A **weakness**.*

Comments

For the most part, eighth graders' performance on math application items fell within the satisfactory range; however, there were problems with calculating averages and solving word problems. In particular, eighth graders had trouble determining what additional data were needed to solve a problem, or which data were extra (Items 14, 15). Areas where students performed satisfactorily include

- Identifying facts in a short word problem,
- Determining the right equation to solve a problem,
- Applying labels (meters, liters) correctly, and
- Calculating perimeter.

BEST COPY AVAILABLE

BEST COPY AVAILABLE

MATH COMPUTATION

The computation portion of the test measures basic arithmetic skills: adding, subtracting, multiplying and dividing whole numbers (positive and negative), fractions, and decimals. Though problems were not intended to be tricky, they were constructed in such a way that common errors would lead students to select an incorrect option; this type of format helps reviewers make an educated guess about the cause of student errors, and perhaps take steps to improve performance.

GRADE 4

Summary of Performance
1 STRENGTH 3 WEAKNESSES

Item Performance

Item 13 Adding Decimals

89.7% correctly solved the problem 3.282 + 0.509. Range was 67.1 to 84.7. *A **strength**.*

Item 20 Multiplying, Regrouping

Only 47.3% correctly solved 375 x 18. Among those who responded incorrectly, many apparently handled the product of the second multiplication (by 1) incorrectly. Range was 59 to 81.3. *A **weakness**.*

Item 21 Division

53.7% correctly divided 1424 by 4: a straight division problem with no remainder. Range was 54.7 to 76.7. *A **weakness**.*

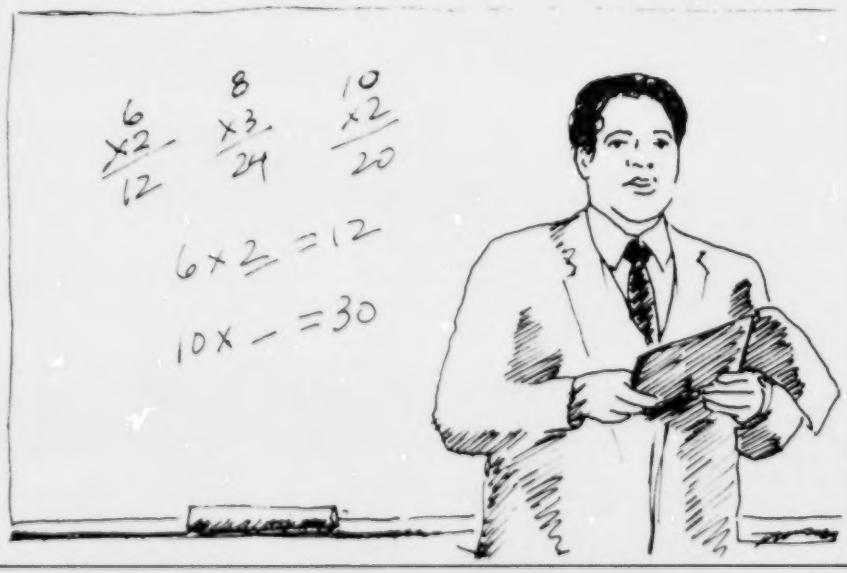
Item 22 Division With Remainder

Not surprisingly, this item, involving division with a remainder, gave students even more trouble than Item 21. Many who answered incorrectly simply got the remainder wrong; others, however, slipped an extra zero into the basic answer. Only 35.5% correctly divided 484 by 31. Range was 48.3 to 72.... *A **weakness**.*

Comments

Fourth graders' performance on addition and subtraction items was generally satisfactory, though only one item (13) in this area revealed a strength. Division, both with and without a remainder, and multiplication with regrouping posed the real problems. Areas where students performed satisfactorily include

- Adding and subtracting whole numbers
- Subtracting decimals
- Multiplying with no regrouping



GRADE 8

Summary of Performance
0 STRENGTHS 3 WEAKNESSES

Item Performance

Item 21 Subtracting Negatives

Only 34.7% could solve 436−(-312). Many students came up with the correct numeral, but mistakenly thought it should be negative rather than positive. Range was 56 to 83. *A **weakness**.*

Item 23 Multiplying Negatives

Students had similar difficulties with Item 23. Only 49.2% could solve (-23) x (-7). Again, many came up with the correct numeral, but thought it should be negative, rather than positive. Range was 59.7 to 86.3. *A **weakness**.*

Item 28 Dividing Fractions

60% correctly divided 3 2/7 by 1/7. This was very close to the satisfactory range—60.7 to 80—but not close enough. *A **weakness**.*

Comments

The most apparent problems occurred with multiplying and subtracting negatives; many students apparently could not determine whether the answer should be a positive or negative number. Satisfactory areas of performance include

- Adding negatives,
- Adding and subtracting mixed numbers and decimals,
- Multiplying and dividing decimals,
- Multiplying fractions, and
- Dividing with a remainder.

In the case of dividing with decimals, however, 69.6% responded correctly, and the minimum level here was 69.0%; in other words, performance was just barely within the satisfactory range. Conversely, on Item 28, notice that student performance just missed falling within the satisfactory range.

MATH CONCEPTS

Math concepts cover such things as number properties, place value, rounding, telling time, using units of measurement appropriately, equivalent fractions, and geometric shapes. These items measure a student's understanding of information presented, but do not require any actual calculation.

GRADE 4

Summary of Performance
3 STRENGTHS 2 WEAKNESSES

Item Performance

Item 23 Words for Number

89.2% of the students knew how to translate the figure 2,047 into words: two thousand forty-seven. Range was 74.3 to 89. ***A strength.***

Item 28 Identifying Even/Odd Numbers

Given four from which to choose, 64.6% correctly identified 582 as an even number. Range was 71.2 to 90.3. ***A weakness.***

Item 34 Feet in a Yard

Only 45% of the fourth graders tested knew how many feet were in a yard. The response "12" attracted many. Range was 60 to 86.7. ***A weakness.***

Item 35 Multiplicative Identity

88.6% knew that 3 x 0 was equal to zero. Range was 65.7 to 85. ***A strength.***

Item 36 Identifying Fractions

Given a series of whole numbers and one fraction, 88.3% correctly identified 3/7 as a fraction. Range was 69 to 87.5. ***A strength.***

Comments

The performance of fourth graders on the concepts subtest produced some surprising results; that they would have difficulty with even/odd and feet in a yard was news to their teachers. Areas of satisfactory performance include

- Translating a fraction into words,
- Determining the number of grams in a kilogram, and
- Picking the right unit to describe length of a fish or weight of an airplane

Incidentally, policies established by the Content Panel call for metric units to be covered in detail on the assessment. It is clear from the Interpretive Panels' ratings and comments, however that metrics are not considered a particularly important part of math instruction in Alaska.

GRADE 8

Summary of Performance
0 STRENGTHS 6 WEAKNESSES

Item Performance

Item 30 Math Laws

Only 27.8% knew that if 3/A = B, then A = 3/B. Panel members expected the item to be difficult, but the performance of students was not much better than guessing level. Range was 38.7 to 64.3. ***A weakness.***

Item 31 Math Laws

When asked to select which one of four equations demonstrating mathematics principles was always correct, only 34.7% correctly chose (2 + A) x 4 = (2 x 4) + (A x 4). Range was 43.7 to 72.3. ***A weakness.***

Item 34 Triangle Properties

36% knew that the hypotenuse of a right triangle is always its longest side. Range was 51.3 to 77.7. ***A weakness.***

Item 35 Sequences

Here students were asked to identify the set of numbers correctly ordered from LEAST to GREATEST. 63.2% identified the right answer: -9, -8, -7, -4, 1. Those who answered incorrectly apparently had difficulty comparing values of negative and positive numbers: e.g., knowing whether -8 was greater or smaller than 1. Range was 64 to 89.2—just a shade above the performance level—but still ***a weakness.***

Item 37 Estimation Checking

In this multiplication of 6,011 by 4,993, students were asked to estimate the answer, not to work the problem (though some may have attempted to work it out). 56.5% correctly identified the answer (30,012,923) by estimating the size of the response and noticing that the number had to end in 3. Range was 62.5 to 88.1. ***A weakness.***

Item 40 Pi

Given four statements about pi, students were asked to identify the one that was NOT TRUE. Only 48.7% knew that the value of pi does NOT depend on the size of the circle. Range was 53.7 to 80.3. ***A weakness.***

Comments

Eighth graders' performance on math concepts reveals six areas of weakness and no strengths. Students did, however, perform satisfactorily on the eight items covering these skills:

- Understanding the properties of circles,
- Understanding the properties of lines,
- Applying fraction terms (e.g., numerator, denominator) correctly,
- Working with greatest common factors and fraction equivalents,
- Identifying a reciprocal,
- Filling in the right number to make an equation true, and
- Rounding a decimal to the nearest tenth.

As in the Computation section, negative numbers posed problems for many students (see Item 35). Equations representing math laws (Items 30 and 31) were troublesome too. Reviewers apparently sensed these items would be difficult; they set the minimum level for Item 30 at only 38.7%—the lowest for any eighth grade math item. Still, performance fell more than 10 percentage points below this minimum—a definite weakness. In addition, almost two-thirds of the students tested had problems with triangle properties. And fewer than half knew that the value of pi never varied.

BEST COPY AVAILABLE

BEST COPY AVAILABLE

READING COMPREHENSION

Reading Comprehension items consist of a reading passage followed by two or three questions based on that passage. The questions are designed to measure students' ability to identify the main idea, distinguish fact from opinion, draw inferences, reach conclusions, determine the order in which major events occurred, understand how ideas are organized, and so on. Great care was taken to ensure that the reading level of each passage was appropriate for the grade level.

GRADE 4

Summary of Performance
5 STRENGTHS 1 WEAKNESS

Item Performance

Item 2 Story: Sequence of Events

Given a list of four events from a story, 90.9% were able to say which event had occurred last. Range was 70 to 90.1. ***A strength.***

Item 4 Story: Main Idea

After reading a short (100-word) story about rivers, 81.9% were able to correctly identify the main idea. Range was 60.3 to 80.7. ***A strength.***

Item 5 Word Meaning

In this item, based on the same passage as Item 4, 53.5% knew that the author used the term "dead river" to mean there were "no living animals in the river." Range was 55.5 to 77.9. ***A weakness.***

Item 6 Facts Paraphrased

In this item, also based on the river story, 85.2% could recognize a specific fact paraphrased from the original. Range was 61.5 to 79.4. ***A strength.***

Item 7 Paragraph Organization

81% could determine that a paragraph on the pony express was organized by "telling how mail delivery had changed across many years." Range was 52.6 to 77. ***A strength.***

Item 12 Inference

After reading a fact-based article on hurricanes, 77.5% correctly inferred that a hurricane would most likely strike on a "windy day with lots of rain." Range was 56.2 to 76.8. ***A strength.***

Comments

Fourth graders gave their best reading performance in this area, with five strengths and only one area of weakness. Areas of satisfactory performance include:

- Recognizing a fact from the story,
- Identifying the conclusion of a story,
- Recognizing the author's opinion,
- Perceiving the author's attempt to persuade the reader, and
- Identifying the most important point.

Fourth graders gave their best performance in this area, with five strengths and only one area of weakness.

GRADE 8

Summary of Performance
3 STRENGTHS 2 WEAKNESSES

Item Performance

Item 4 Facts in a Passage

In an item based on a brief account of Mark Twain's life, students were asked which of four statements was NOT a fact. 58.4% correctly identified this opinion: "Twain is the best of all American authors." Range was 65.9 to 86.7. ***A weakness.***

Item 8 Interpreting Information from a Poem

Items 8 and 9 were based on Robert Frost's poem "Provide, Provide." Asked which of four statements Frost would likely agree with, 71% correctly said, "Fame does not make up for a lack of friends and money when a person is old." Range was 47.9 to 65. ***A strength.***

Item 9 Author's Purpose

Here students were asked to determine Frost's purpose in writing "Provide, Provide." 69.3% correctly responded "To advise people to prepare for old age and death." Range was 40.4 to 57.5—well below the students' actual performance. ***A strength.***

Item 10 Support for Statements

Given a paragraph on road maps provided by AAA, only 29.7% recognized that the author's claim to have the "only accurate" map around could not be supported by the facts presented. Range was 51.2 to 75.9. ***A weakness.***

Item 15 Drawing Inferences

After reading a brief account of Sacajawea's life, 77.7% inferred that the statement "Sacajawea was brave and adventurous" was likely to be true, even though it was not stated in the story. Range was 57.1 to 76.2. ***A strength.***

Comments

Eighth graders revealed one strength in drawing an inference (Item 15), and two areas of strength (see Items 8 and 9) in their analysis of Frost's poem "Provide, Provide;" also, their performance was satisfactory on a third item about the poem. It is interesting to note that range levels were set relatively low for these three poetry items in comparison to levels for other eighth grade reading items. Eighth graders also performed satisfactorily in other areas, including

- Recognizing the main idea of a story,
- Understanding a figure of speech,
- Correctly matching pronoun and referent,
- Recognizing exaggeration, and
- Identifying an author's method of organization.

The weakness on Item 10 (support for statements) represented the single biggest discrepancy between actual performance and Minimum level on the entire eighth grade reading test—more than 20 percentage points.

BEST COPY AVAILABLE

STUDY SKILLS

Can students use a dictionary effectively? How about a map, index, graph, chart, or table? This section of the test measures students' ability to find and selectively use needed information. Readers should take special care in interpreting results from this section, keeping in mind that some study skills covered in the assessment test may not be part of some Alaska districts' curriculum.

GRADE 4

Summary of Performance
1 STRENGTH 8 WEAKNESSES

Item Performance

Item 16 Use of the Dictionary

Given a portion of an actual dictionary page, 77.5% determined that this statement was NOT true: "A book would be easy to MOLD." Range was 53.8 to 75.3. ***A strength.***

Item 17 Use of the Dictionary

Surprisingly, performance on this item—parallel in form to Item 16—was rather low. Only 44.9% recognized this statement as UNTRUE: "Bread is more likely to MOLD in dry places than wet places." The range was set relatively low for this item—45.6 to 69.1—but not low enough. ***A weakness.***

Item 18 Dictionary Guide Words

Here again, many fourth graders had trouble. Given the guide words "molasses" and "Monday," only 45.1% could tell that "monastery" would appear on the same page. Range was 51.8 to 70.3. ***A weakness.***

Item 19 Using an Index

Items on using an index proved even more troublesome for students. Only 42.6% could use an index to locate the right pages for a specific topic. Range was 70.8 to 90.6. Panelists apparently did not consider this a difficult item, and performance fell almost 30 percentage points below expectations: the largest discrepancy on the fourth grade reading test. ***A weakness.***

Item 20 Using an Index

Performance on this item, based on the same index, was somewhat better, but still well below expectations. 58.2% could identify the BEST place to look for specific information. Range was 70.9 to 90.9 (almost identical with Item 19). ***A weakness.***

Item 21 Using an Encyclopedia

Only 50% knew that "farming" would be covered in an encyclopedia with guide letters ENI to FIN. Range was 63.5 to 84.1. ***A weakness.***

Item 22 Where to Look for Information

58.6% knew that the best place to look for definitions used in a book would be in the book's glossary. Range was 61.5 to 84.7. ***A weakness.***

Item 26 Following Directions

49.8% could follow directions for making a simple drawing, then answer a question based on the results. Range was 51.5 to 73.8. Not a big difference here, but still ***a weakness.***

Item 28 Using Table of Contents

Only 59.3% could use a table of contents to find information on a specific topic. Interestingly, in Item 27, based on the same table of contents, 80.1% had responded correctly to a parallel item. That performance was well within the satisfactory range—close to indicating a strength. But on Item 28, students had to do a little more reading to get to the right answer. Range was 62.6 to 85.6. ***A weakness.***

Comments

Study Skills proved a difficult area for many fourth graders. Performance was strong on only one item: 16 (using the dictionary). But it fell far below expectations on several other items—those dealing with the use of an index (Items 19, 20), encyclopedia (21, 22), and table of contents (28). Student performance was satisfactory on the 10 items covering these skills; items covering these skills:

- Selecting the best encyclopedia heading to reference for specific information,
- Using a map key,
- Reading a line graph,
- Using a table of contents,
- Using a glossary, and
- Reading a road map.

BEST COPY AVAILABLE

STUDY SKILLS

continued

GRADE 8

Summary of Performance
4 STRENGTHS 3 WEAKNESSES

Item Performance

Item 19 Using Map Key

93.5% could use a key to tell which point along a race course offered both water and first aid. This was the single best performance by eighth graders on the reading test. Range was 67.1 to 88.7. ***A strength.***

Item 20 Reading a Map

84.8% could use a map of projected spacecraft routes to determine proximity of Voyager 2 to specific planets on specific dates. Range was 58.7 to 79.6. ***A strength.***

Item 21 Out-of-Date Books

Given four book titles and years of publication, only 38.2% could tell that *What You Should Know About Investing Your Money*, published in 1951, would contain the most out-of-date information. Range was 52.1 to 74.2. ***A weakness.***

Item 22 Biased Statements

Item 22 was similar in format to 21—and again, many students had trouble. Given four quotations and their sources, only 33.8% could determine which statement would be most credible. Range was 45.9 to 65.4. ***A weakness.***

Item 23 Using a Chart

74.4% could use a chart to determine which information about various campgrounds was most likely to be correct. Range was 52.1 to 70.1. ***A strength.***

Item 24 Following Directions

For this item, students had to change position on a small graph by following a set of five short directions, then correctly identify their final position. 78.6% were successful. Range was 55 to 77.5. ***A strength.***

Item 25 Using an Index

Recall that in the fourth grade test, items on the dictionary (15, 16) and on tables of contents (27, 28)—though apparently similar in structure and format—produced very different levels of student performance. This also proved true on the eighth grade test with Items 25 and 26, both based on the same section of an index. While performance on Item 26 fell well within the satisfactory range (more than 10 points above Minimum), on Item 25, only 59.2% could use an index to determine where to look for specific information. Range was 68.3 to 90.4. ***A weakness.***

Comments

In Study Skills, as in Comprehension, eighth graders' strengths outnumbered weaknesses. They had the greatest difficulty with Items 21 and 22, both calling for judgments on the information presented. Interestingly, performance was well within the satisfactory range (64.4% responding correctly, with a range of 46.7 to 67.5) on Item 29, which presented four quotations and their sources, then asked students which statement was "most believable." This item was nearly identical structurally to Item 22, which asked students which of four statements was most "unbiased and objective"—yet note the differences in performance.

There were no real "borderline" performances in the satisfactory range; that is, in most cases, satisfactory performance was a good 5 to 10 percentage points (or more) above the Minimum level. Eighth graders gave satisfactory performance on these skills:

- Using a chart, map, or graph to obtain information,
- Using an index,
- Using a dictionary, and
- Determining which of four sources was "most believable."

WORD SKILLS

This section of the test covered skills many Alaska educators consider vital to reading proficiency: knowledge of regular and irregular vowel sounds, consonant combinations, possessives, prefixes and suffixes, synonyms and antonyms, and multiple meaning words.

GRADE 4

Summary of Performance
0 STRENGTHS 8 WEAKNESSES

Item Performance

Item 31 Beginning Sounds

Only 61.5% recognized that "drum" and "drop" have the same initial blend sound. Students often chose a rhyming word or a word with only the initial "d" sound instead. Range was 82.6 to 96.3. *A **weakness**.*

Item 32 Beginning Sounds

And only 60.8% recognized that "chair" and "check" have the same initial blend sound. Range was 82.3 to 96.5. (Note how high the range levels were on Items 31 and 32.) *A **weakness**.*

Item 34 Vowel Sounds

75.5% recognized that "me" and "need" have the same vowel sound. Range was 80 to 94.1. *A **weakness**.*

Item 35 Vowel Sounds

Performance was about the same here as on Item 34; 74.5% knew that "plow" and "round" have the same vowel sound. Range was 76.4 to 91.8. *A **weakness**.*

Item 36 Vowel Sounds

Again, similar performance levels: 75.5% knew that "gate" and "aid" have the same vowel sound. Range was 77.9 to 92.1. *A **weakness**.*

Item 37 Vowel Sounds

On this item the Minimum level dropped considerably—but unfortunately, so did performance. Only 56.1% recognized that "cross" and "hall" have the same vowel sound. Range was 66.8 to 82.2. *A **weakness**.*

Item 38 Plurals

57.7% recognized that the plural "horses" was the right word to complete this sentence: "Eight _____ were hitched to each covered wagon." Range was 62.5 to 82.3. *A **weakness**.*

Item 39 Irregular Plurals

59.6% knew that "deer" was the correct word to complete this sentence: "Several _____ crossed the road." Range (like the performance level) was similar to that for Item 38—62.8 to 83.2. *A **weakness**.*

Comments

This seemed the single most troublesome section of the Reading test for fourth graders—the only section in which weaknesses actually outnumbered satisfactory performances—8 to 7. There were no areas of identified strength at all. Word Skills is the subtest in which the Interpretive Panel members most consistently overestimated the fourth grade students' performance. They clearly thought the items rather easy—setting Minimums at over 80% for the consonant items. Performance on both items, unfortunately, fell more than 20 points below those Minimum levels.

Vowel sounds proved troublesome as well. Out of five items on vowel sounds, students performed satisfactorily on only one—and that was a close call, with performance level at 80.8%, and Minimum at 80.3.

Word Skills is the subtest in which the Interpretive Panel members most consistently overestimated the fourth grade students' performance.

Students did perform satisfactorily on Word Skill items measuring these skills:

- Use of present and past tenses
- Identifying "weren't" as a contraction for "were not"
- Recognizing "near" as a synonym for "close"
- Identifying "rude" as an antonym for "polite"
- Inferring the correct meaning of "checked" from sentence context: "The doctor checked her weight."

GRADE 8

Summary of Performance
3 STRENGTHS 2 WEAKNESSES

Item Performance

Item 32 Vowel Sounds

87.4% recognized that "priest" and "teach" have the same vowel sound. Range was 61.7 to 85. *A **strength**.*

Only 33.8% knew that the plural possessive boys' was the correct word to complete this sentence: All the _____ names were called.

Item 34 Plural Possessives

Only 33.8% knew that the plural possessive boys' was the correct word to complete this sentence: "All the _____ names were called." Range was 58.3 to 84.2. *A **weakness**.*

Item 36 Past Tense

89.7% recognized that "lied" was the correct form to complete this sentence: "Roger _____ about his age." This was almost a borderline case, with the satisfactory range running 68.3 to 89.2. *A **strength**.*

Item 37 Suffix

Given four words with the same root but different suffixes, 87.7% correctly chose "nutritionist" as the equivalent of "expert in nutrition." Range was 61.2 to 81.2. *A **strength**.*

Item 38 Prefixes

Given four words with the same root, only 33.3% recognized "enthroned" as the equivalent of "put on the throne." Notice the striking difference in performance between this item and the similar Item 37. Range here was 51.7 to 75.4—lower than that for Item 37, but not low enough. *A **weakness**.*

BEST COPY AVAILABLE

WORD SKILLS

continued

Comments

The ratio of strengths to weaknesses was remarkably similar on all sections of the eighth grade Reading test. Here again, strengths outweighed weaknesses (just barely)—this time by 3 to 2. In all cases where a strength was indicated, however, student performance was just a shade above the Desired level. But in Items 34 (Plural Possessives) and 38 (Prefixes), the actual performance fell far below Minimum expectations.

The performance of the grade 8 students was satisfactory on items measuring these skills:

- Recognizing that "pouch" has the same vowel sound as "crown"
- Recognizing that "deep" has the same vowel sound as "thief"
- Knowing that "gentle" has the same ending sound as "metal"
- Using the present tense "watches" correctly to complete a sentence
- Recognizing "moved" as a synonym for "transported"
- Recognizing "denied" as an antonym for "admitted"
- Inferring the correct meanings of "covered" and "pool" from context: "The speaker *covered* all the important points." "They will *pool* their money and solve their problem."

In all cases but one, satisfactory performance was more than 10 percentage points above the Minimum level.

HOTLINE!!!!

We can't really keep calling it a HOTLINE unless we hear from you. Do you have a question, comment, or suggestion about the Alaska Statewide Assessment? Before you lose that thought, call COLLECT (503) 223-3396. Someone knowledgeable about the Assessment will be happy to talk with you.

USING THE RESULTS IN ALASKA DISTRICTS

It is critical to recognize that the interpretations presented in this newsletter are based on a statewide perspective; interpretations of student performance district by district could differ markedly. In other words, suppose District X provided no direct instruction in a given skill—yet 50% of the District X students responded correctly to an item measuring that skill. This performance might seem, from the district perspective, extremely positive—even if that 50% fell below the established statewide Minimum level.

How then might individual districts look at these results? First, most will want to carefully consider the findings of the Interpretive Panels, noting identified strengths and weaknesses statewide, and compare those findings to what results show about local performance:

- Are strengths and weaknesses similar?
- Are there striking differences? (e.g., an identified weakness statewide may emerge as a strength for some districts—or vice versa)
- How close to "cutoff" levels are strengths and weaknesses? That is, a weakness right on the borderline—say 1% below the established Minimum—will usually be less cause for concern than one, say, 20% below the Minimum.

Second, local educators will want to ask, "Do the state priorities match local priorities?" In some cases, they may not. For example, consider the grade 8 Reading test. Here the Interpretive Panel set relatively low Minimum and Desired levels on poetry items (7, 8, and 9)—and relatively high levels on several Study Skills items (19, 25, 26). Such perspectives may or may not reflect individual districts' priorities. The point is, levels set by the Interpretive Panels are only guidelines—not inflexible criteria for judging *all* performance.

Finally, based on a review of actual student performance, districts may wish to further pinpoint identified strengths and weaknesses. Remember that while the Alaska Assessment test results can indicate broad areas of strength or weakness, further research and testing would be of assistance in developing a sound base for instructional or curricular modification.

HAS STUDENT PERFORMANCE IMPROVED SINCE 1981?

The tables on this page show, subtest by subtest, how students performed in 1981 compared to 1983. In making comparisons, please keep in mind that the 1981 and '83 tests were highly similar—but NOT identical. Items for '83 were modeled on those for '81 so that some degree of comparison would be possible. They were similar in format, content, and reading level, and the same numbers of items were included on each subtest.

Small differences may be due simply to item variations rather than an actual change in performance of students.

But even items which are carefully constructed to measure the same skill have slight variations in difficulty. The important point to remember while examining the information below is that although comparisons can and should be made across years, small differences may be due simply to item variations rather than an actual change in the performance of students.

The Interpretive Panels – An Important Variable

The Desired and Minimum levels set by the Interpretive Panels for 1981 and 1983 were slightly different, of course. For example, take Item 4 on the grade 4 Reading test: identifying the main idea of a reading passage. Here the Desired level for 1981 was 84%—for 1983, 80.7%.

Such differences are hardly surprising, considering that the Panel members were different for the two years, and that the 1983 Panels made absolutely no attempt to match levels for 1981 (nor would it have been appropriate or desirable to do so, given inevitable shifts in curricular priorities and instructional goals). What is surprising is that in most cases differences were relatively slight, although levels tended to be slightly lower on average for '83. In many cases, in fact, Minimum and Desired levels for '81 and '83 on comparable items were only two or three percentage points apart.

This means that, generally speaking, the 1981 and 1983 Interpretive Panels had similar expectations for students, and also made similar judgments about relative item difficulty. At the same time, keep in mind that very slight differences in performance can make the difference between a strength and satisfactory performance—or between satisfactory performance and a weakness. Recognizing that this is so, reviewers will want to pay special attention to borderline performances—those just barely within the range of one rating or another.

If the strengths and weaknesses obtained in 1981 are compared to 1983, the performance on the 1983 test looks slightly better. Some portion of this improvement is due to higher scores on the tests, the rest due to slightly lower expectations from the 1983 Interpretive Panel. Although the composition of the two panels was quite similar, the '83 Interpretive Panel set lower expectations in 66 cases, higher in only 18. If the performance levels set by the 1981 Interpretive Panel were applied to the 1983 results, the performance of the students would appear slightly lower. In short, it appears that the expectations of Alaska educators are slightly lower than they were in 1981. But the difference is small enough that it could be due to differences in the composition of the panels, rather than an overall trend.

1981 Interpretive Panel Results

	Mathematics			Reading		
	Application	Computation	Concepts	Comprehension	Study Skills	Word Skills
GRADE 4						
Strength	3	2	2	3	0	0
Satisfactory	6	7	8	11	1	3
Weakness	1	3	5	1	14	12
Total Items	10	12	15	15	15	15
GRADE 8						
Strength	0	0	0	0	0	2
Satisfactory	8	5	7	12	8	7
Weakness	8	7	8	3	6	4
Total Items	16	12	15	15	14	13

1983 Interpretive Panel Results

	Mathematics			Reading		
	Application	Computation	Concepts	Comprehension	Study Skills	Word Skills
GRADE 4						
Strength	3	1	3	5	1	0
Satisfactory	4	8	10	9	6	7
Weakness	3	3	2	1	8	8
Total Items	10	12	15	15	15	15
GRADE 8						
Strength	1	0	0	3	4	3
Satisfactory	10	9	9	10	7	8
Weakness	5	3	6	2	3	2
Total Items	16	12	15	15	14	13

EDUCATORS' EXPECTATIONS CONTINUE HIGH

Results from the 1983 Assessment Interpretation

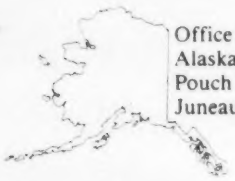
When I joined the Alaska Department of Education in 1976, the state had already begun to measure students' achievement level on a statewide basis through its Alaska Assessment Tests. Those first years of the assessment produced some high quality tests and a testing process that functioned fairly well. The results were interesting indications of just how much Alaska's students knew about some critical basic skills objectives. But something was missing from those early tests.

When test results became available, they showed what percentage of students answered each item correctly. But there was no way to know if the results showed a good or bad performance. During the first assessments, we had only the student scores and thus couldn't compare the performance against the expectations held for students by their teachers.

This shortcoming was corrected for the last two assessments (1981 and 1983) through the appointment of the Assessment Interpretive Panels. These panels examine each of the assessment items and establish performance expectations. When these expectations are later compared to the actual results, a set of strengths and weaknesses of the students' performance results. These strengths and weaknesses are the best indication we have of how well Alaska students are learning the things we expect them to learn. This Assessment Report discusses the strengths and weaknesses shown by the 1983 testing; I hope you will consider them carefully.

Al Hazelton
Supervisor of Research

These strengths and weaknesses are the best indication we have of how well Alaska students are learning the things we expect them to learn.



Office of Research & Evaluation
Alaska Department of Education
Pouch F
Juneau, AK 99811

PRESORTED
FIRST CLASS
MAIL
U.S. POSTAGE
PAID
JUNEAU, AK
PERMIT NO. 224

ASSESSMENT REPORT—Please Circulate To All Interested Individuals

ASSESSMENT REPORT 9: An Update on the Alaska Statewide Testing Program

What Alaska Students Should Learn: Perspectives From Those Who Teach Them

In March 1983, Alaska administered its fourth Statewide Assessment in Math and Reading. Regular readers of this newsletter have seen the results, long since tabulated and returned to schools and districts. And an analysis of those results, the product of an in-depth review by an Interpretive Panel, formed the basis for our last report.

This update looks at the assessment results in light of recommendations made by Alaska teachers' professional associations, not as part of the assessment process *per se*, but in an effort to define the reading and math content most important for Alaska students. Their recommendations offer a challenge to educators, parents and citizens concerned with quality education to review current local objectives, and to consider whether present curriculum offers students every chance to develop critical skills.

ALASKA'S IMPORTANT LEARNING OUTCOMES

During the late spring of 1983, five professional teachers' associations were commissioned by the Alaska Department of Education to develop position papers on what they believed were important learning outcomes for Alaska students in grades 6 and 12. Two of the five papers are particularly relevant to the Alaska Statewide Assessment:

- Mathematics in Alaska Schools — A position paper by the Alaska Council of Teachers of Mathematics (ACTM)
- Reading in Alaska Schools — A position paper by the Alaska State Reading Association (ASRA)

The goals and objectives outlined by ACTM and ASRA are, in many cases, closely related to the content covered by the reading and math Assessment.

Although prepared long after the 1983 test content was established by the statewide Content Panels, the goals and objectives outlined by ACTM and ASRA are, in many cases, closely related to the content covered by the statewide assessment. This report examines test results in light of relevant expectations presented in these two position papers.

TM 850 546 (2 of 3)

Of course, it is vital to keep in mind that the objectives established by ACTM and ASRA are intended for students in grades 6 and 12. Since the Alaska Assessment tests are administered at grades 4 and 8, great care must be exercised in comparing student performance against expectations that may be considerably above or below grade level. For this reason, the objectives for grade 12 are not included in the comparisons that follow.

Those interested in reviewing the entire set of position papers—which cover science, social studies and English as well as math and reading—should request a copy of "What Alaska Students Should Learn: Perspectives from Those Who Teach Them" from the Alaska Department of Education, Pouch F, Juneau, Alaska 99811.

REVIEWING THE INTERPRETIVE PANEL PROCESS

Assessment Report 8, issued a little over a month ago, summarized the Interpretive Panel's review of student performance on the 1983 Alaska Statewide Assessment tests. These tests, designed to measure performance of Alaska fourth and eighth graders in reading and math, comprise six subtests:

READING
- Comprehension
- Study Skills
- Word Skills

MATH
- Application
- Concepts
- Computation

The Interpretive Panel members—more than 70 classroom teachers and administrators who are specialists in the content areas of math and reading—were nominated by district superintendents and members of the Assessment Content Panels, the educators who help design and structure the assessment tests. Last fall, these specialists carefully reviewed every item on the 1983 tests and set levels of expectation for student performance. Assessment program staff then compared actual performance to those levels to identify salient strengths and weaknesses.

Strengths and weaknesses are determined this way. Before seeing any actual results, Interpretive Panel members set two levels of expectation:
- First, a Minimum level, representing the lowest percentage of students answering correctly that Panelists would regard as acceptable;
- Second, a Desired level, representing the percentage of students that Panelists would like to see answer correctly.

Suppose that on a given item the Minimum level were set at 70 percent, and the Desired level at 85 percent. Then if fewer than 70 percent of the students answered correctly, performance would be said to indicate a WEAKNESS. If somewhere between 70 and 85 percent answered correctly, performance would be judged SATISFACTORY. And if more than 85 percent answered correctly, performance would be said to show a STRENGTH.

Interpretive Panelists consider the relative difficulty of each item, the importance of the skill being measured, and appropriate expectations and goals for students at each grade level.

The perspectives of the Interpretive Panel are significant because they represent, in effect, not only a judgment of student performance, but also a critical review of the test items themselves. In making their judgments, Interpretive Panelists consider the relative difficulty of each item, the importance of the skill being measured (relative to other skills covered by the test), and appropriate expectations and goals for students at each grade level. Recall that Interpretive Panel members set their expectation levels **without** seeing actual performance results.

The performance levels set by the Interpretive Panel are useful not only to those who want a guideline for judging student performance, but also to those responsible for modifying and refining the Alaska Assessment. For example, extremely low Minimum levels might indicate that an item was considered too difficult for grade level; conversely, extremely high Minimums might indicate that an item was considered too easy for grade level. Such judgments help direct the review of test items that precedes each assessment.

THE IMPORTANCE OF OTHER PERSPECTIVES

When developing the Assessment tests, the Content Panel members make every effort to be aware of and sensitive to the objectives held by individual districts; yet local expectations may justifiably differ slightly from those set by the Panelists. Why? Remember that Alaska has no statewide curriculum. Each of the state's 53 districts independently determines what content will be taught at a given grade level—and when. Thus, not every district objective will be reflected in a statewide assessment.

There is also the possibility that objectives that are important to an individual district weren't selected for statewide testing. In short, although the Alaska Assessment objectives are well matched to most districts' objectives, there may be slight differences from district to district. For that reason, it is instructive to consider alternate sets of objectives when reviewing results from the Statewide Assessment.

How closely do the objectives identified by the Content Panels—composed of teachers and administrators from around the state—reflect what Alaska professional organizations think is important today? Following is a summary of the recommendations made by the statewide reading and mathematics professional associations that are most closely related to the Assessment test, together with a report of student performance on relevant items. Notice that while these recommendations were established totally independently of the Assessment test, there is frequently a striking match between the professional associations' recommendations and test content.

COMPARING MATHEMATICS OUTCOMES TO ACTM RECOMMENDATIONS

Let's take a look at student performance in light of the recommendations issued in Spring 1983 by the Alaska Council of Teachers of Mathematics. While the following list of recommendations is not comprehensive (the complete ACTM list is somewhat more detailed), it is still an instructive activity. Our purpose here is to focus on those areas (and there were many) specifically covered by the statewide assessment in order to help readers consider test results from the perspective of what Alaska's own math teachers and specialists have said was most important.

The ACTM divided their recommendations for the elementary level (grade 6) into six categories:

- Computation,
- Problem Solving,
- Measurement,
- Geometry,
- Graphing, and
- Computer Literacy.

Numerous concepts within the first four categories are covered by the Statewide Assessment test.

Readers should keep in mind that the skills list was developed specifically for grade 6 (ACTM also developed a separate list for grade 12), whereas the assessment is administered at grades 4 and 8. Some flexibility is essential, therefore, in comparing actual performance to ACTM expectations. (In the charts below, numbers in parentheses indicate the relevant items from the 1983 Assessment.)

Comparison Summary

ACTM Recommendations	Grade 4 Performance	Grade 8 Performance
Recommendation 1–Students should be able to read, write, and order whole numbers, fractions, and decimals.	Strengths on reading numbers (item 23) and identifying a fraction (36). Satisfactory on ordering whole numbers (24).	A weakness on the one relevant item (35), ordering whole numbers.
Recommendation 2–Students should be able to give the place value of a digit in standard numerals and decimals.	Satisfactory on identifying place value (25) and two rounding items (26, 27).	No matching items.
Recommendation 3–Students should be able to add, subtract, multiply, and divide using whole numbers.	Satisfactory on 5 items (11, 12, 14, 15, 16) requiring adding and subtracting whole numbers and 2 items on multiplying without regrouping (18, 19). Weaknesses on multiplying with regrouping (20) and dividing whole numbers (21, 22).	Satisfactory on adding negative whole numbers (17) and dividing whole numbers with remainder (27). Weaknesses on subtracting negative whole number from positive whole number (21) and multiplying negative whole numbers (23).
Recommendation 4–Students should be able to choose the correct algorithm (process) to carry out a computation.	Satisfactory on 2 word problems (3, 4) requiring addition and multiplication, respectively. Weakness on a word problem involving subtraction (9).	A strength on 1 word problem requiring addition (3), but satisfactory on another similar item (10). Satisfactory on a word problem requiring multiplication (5). Weaknesses on word problems requiring subtraction (12) and multiplication of decimals (16).
Recommendation 5–Students should be able to add, subtract, multiply, and divide with decimals and fractions at an introductory level, and should begin to understand the relevance of fractions to decimals.	A strength on adding decimals (13), but satisfactory on subtracting decimals (17).	Ten satisfactory performances, as follows: adding decimals (18), subtracting decimals (22), multiplying decimals (24), dividing decimals (25), subtracting mixed numbers (19), multiplying fractions (26), using fraction terms (36), identifying equivalent fractions (39), identifying a reciprocal (41), and rounding decimals (43). A weakness on dividing fractions (28).

COMPARING MATHEMATICS OUTCOMES TO ACTM RECOMMENDATIONS

Comparison Summary, continued

ACTM Recommendations	Grade 4 Performance	Grade 8 Performance
Recommendation 6—Students should show understanding of a problem, be able to identify the correct process for solving a problem, and systematically apply a set of logical steps to resolve a mathematical problem from everyday life.	Strengths on understanding the problem (2, 7). Satisfactory on word problems involving addition (1), noting exact facts (5), and addition and division (6). A weakness on another word problem requiring noting exact facts (10).	Satisfactory performances on word problems requiring addition and subtraction (1), noting exact facts (4), and understanding the problem (11, 13). Weaknesses on word problems with averages (2), missing data (14) and extra data (15).
Recommendation 7—Students should have knowledge of linear measurement, either standard or metric, for distance, area, and perimeter; and they should be able to calculate weight (mass) and liquid capacity.	A strength on understanding a problem involving area (7). Satisfactory performances on equation for perimeter (4), noting exact facts (5), metric units (31), units for weight (32) and grams in kilogram (33). Weaknesses for equation for area (8) and feet in a yard (34).	Four satisfactory performances: recognizing meter as correct label (6), recognizing correct unit for volume (7), recognizing milliliter as correct label (8), and knowing the definition for perimeter (9).
Recommendation 8—Students should be able to tell time to the second.	Satisfactory on translating time (numerals) into words (29) and telling time from clock face (30).	No matching items.
Recommendation 9—Students should have knowledge of such concepts as point, line, perpendicular, parallel, and congruent.	No matching items.	Satisfactory on line properties (33).
Recommendation 10—Students should know the properties of simple geometric shapes in a two-dimensional plane.	No matching items.	Satisfactory on circle properties (32). Weaknesses on triangle properties (34) and characteristics of pi (40).
TOTALS	**6 Strengths, 7 Weaknesses, 24 Satisfactory**	**1 Strength, 11 Weaknesses, 24 Satisfactory**

Summary Note—

In light of ACTM recommendations, Alaska educators may want to pay particular attention to these areas of identified weakness:

Grade 4
- Solving mathematical word problems
- Determining how to solve a word problem
- Multiplying with regrouping
- Dividing both with and without remainders
- Application of standard measures (for example, feet in a yard)

Grade 8
- Solving mathematical word problems
- Determining how to solve the problem
- Subtracting and multiplying negative numbers
- Dividing fractions
- Numeric sequences, especially involving negative numbers
- Triangle properties
- Characteristics of pi

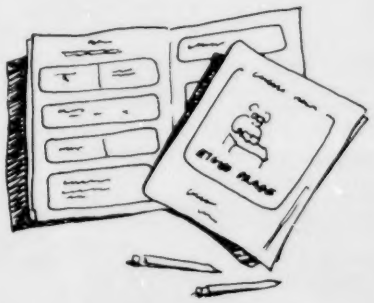
COMPARING READING OUTCOMES TO ASRA RECOMMENDATIONS

Now it's appropriate to consider students' reading performance in comparison to the recommendations issued in Spring 1983 by the Alaska State Reading Association (ASRA). As for math, keep in mind that this is NOT the complete ASRA list. We will focus on those recommendations most relevant to the content of the Alaska Assessment test. Remember also that these ASRA recommendations were directed at students completing grade 6, so that some may seem advanced for grade 4 or slightly below level for grade 8. (A separate list was developed for grade 12.)

The Alaska State Reading Association divided their recommendations into three categories:

- Comprehension
- Study and Survival Skills
- Vocabulary

These three categories are roughly comparable to the three reading assessment subtests: Comprehension, Study Skills and Word Skills.



Comparison Summary

ASRA Recommendations	Grade 4 Performance	Grade 8 Performance
Recommendation 1 – Students will identify main ideas and details at the literal level in information, content, fiction and survival materials.	Strengths in identifying the main idea (item 4) and paraphrasing facts (6). Satisfactory performances in identifying details from the story (1, 13), identifying the main idea (10), understanding use of referent (14), and identifying supporting facts (15).	Four satisfactory performances: identifying the main idea (1), recognizing pronoun referent (3), identifying most important point (11), and identifying details from the story (13).
Recommendation 2 – Students will be able to sequence events in reading and content material.	Two strengths: putting events in sequence (2) and identifying method of organization (7).	Satisfactory in identifying method of organization (6) and putting events in sequence (14).
Recommendation 3 – Students will apply reading skills beyond the literal meaning (i.e., the interpretive and critical/applicative levels).	No matching items.	Strengths in two items on poetry interpretation (8, 9), satisfactory on a third similar item (7). Satisfactory on figure of speech (2).
Recommendation 4 – Students will demonstrate ability to understand material in an oral and written mode, and to draw conclusions and make judgments based on information received.	Strength in drawing an inference (12). Satisfactory in drawing a conclusion (3), identifying an opinion (8), recognizing author's attempt to persuade (9), and drawing an inference (11).	A strength in drawing an inference (15). Satisfactory on recognizing exaggeration (5), judgment based on evidence (12), and identifying most believable source (29). Weaknesses on judgment based on evidence (10, 21) and recognizing bias (22).
Recommendation 5 – Students will demonstrate a basic use of textbooks, reference books, and library resources.	Satisfactory in using an encyclopedia (23), table of contents (27) and glossary (29). But weaknesses in using an index (19, 20), encyclopedia (21) and table of contents (28), and knowing where to find information (22).	Satisfactory on one item about using an index (26), but a weakness on another similar item (25).

COMPARING READING OUTCOMES TO ASRA RECOMMENDATIONS

Comparison Summary, continued

ASRA Recommendations	Grade 4 Performance	Grade 8 Performance
Recommendation 6 – Students will demonstrate use of basic schedules, charts, and directories.	Satisfactory in using a map key (24), reading a chart (25) and reading a road map (30).	Three strengths: using a map key (19), reading a map (20), and reading a chart (23). Satisfactory on reading a chart (16), using a map key (17), and reading a graph (18).
Recommendation 7 – Students will be able to follow oral and written directions.	A weakness in following written directions (26).	Strength on following written directions (24).
Recommendation 8 – Students will develop independent word attack skills to decode unfamiliar words in information, content, and survival reading materials.	Satisfactory in vowel sound (33). Weaknesses in consonant blends (31, 32) and other vowel sound items (34, 35, 36, 37).	Strength on one vowel sound item (32), but satisfactory on two others (30, 31). Satisfactory on ending consonants (33).
Recommendation 9 – Students will understand the structure of words (that is, prefixes and suffixes, root words and endings, and compound words).	Satisfactory in present tense verbs (39), past tense verbs (40), and contractions (42). Weaknesses in noun plurals (38) and irregular plurals (41).	Strengths on past tense verbs (36) and suffixes (37). Satisfactory on present tense verbs (35). Weaknesses on plural possessives (34) and prefixes (38).
Recommendation 10 – Students will understand word relationships (that is, synonyms, antonyms, and homonyms).	Satisfactory performances on items about synonyms (43), antonyms (44), and multiple meaning words (45).	Four satisfactory performances: synonyms (39), antonyms (40), and multiple meaning words (41, 42).
Recommendation 11 – Students will use the dictionary to learn the meaning of unknown words.	A strength on one item about using the dictionary (16), but weaknesses on two others measuring the same thing (17, 18).	Satisfactory on use of the dictionary (27, 28).
TOTALS	**6 Strengths, 16 Weaknesses, 22 Satisfactory**	**10 Strengths, 6 Weaknesses, 25 Satisfactory**

Summary Note—

In light of ASRA recommendations, Alaska educators may want to pay particular attention to the following identified weaknesses:

Grade 4
- Using an index
- Using an encyclopedia
- Using a table of contents
- Knowing where to find information
- Following written directions
- Consonant blends
- Vowel sounds
- Regular and irregular noun plurals
- Use of the dictionary

Grade 8
- Making a judgment based on evidence
- Recognizing bias
- Using an index
- Plural possessives
- Prefixes

1985 ASSESSMENT SCHEDULED

Educators in Alaska's schools and district offices should mark their calendars for the week of March 25-29, 1985. These are the dates of the fifth Alaska Statewide Assessment. During this week, every fourth and eighth grade student in a regular classroom will be tested on important reading and math skills.

Many educators who worked in Alaska during 1977, 1979, 1981 or 1983 probably recall similar testing activities. Administered biennially, the Statewide Assessment was established in the mid 1970s in response to both public and professional needs for information about students' academic performance. Currently, the Assessment provides the only available indication of achievement for groups of students on a statewide basis.

Important Information about the 1985 Testing

While the testing itself will occur during the week of March 25-29, staff in Alaska's districts and schools start to plan for the assessment much earlier than that. Here are seven important points about the 1985 testing.

1. Dates

Testing will take place during the week of March 25-29, 1985. This week best balances the need to test at the end of the school year and the desire to return results to the schools before the summer recess. Local staff should make sure that these dates will not conflict with other district testing sessions.

2. Who Is Tested

With only a few exceptions related to handicapped students, every fourth and eighth grade student will be tested. Testing every student provides schools and districts with complete and continuous indications of performance, something not possible with sampling.

3. Administration

Tests will be administered by regular classroom teachers. In schools where the eighth grade students have multiple teachers, districts should determine the most efficient method of administration. Teachers will not need special training to administer the tests. The directions for the test are straightforward and there are no oral items.

4. Time Requirements

Tests will require two administration periods (one for math and one for reading) of slightly less than one hour each. The mathematics section is usually given first. A minimum of half a day should be allowed between the sections.

5. Test Structure

The assessment test has three sections on mathematics (application, computation and concepts) and three sections on reading (comprehension, study skills and word skills). Each subtest has between 10 and 16 items.

Testing will take place during the week of March 25-29, 1985.

6. Scoring and Reporting

Separate answer sheets will be mailed to the scoring agency by each school (or, in some cases, by the district). Test scores will be reported for each student tested. The scores are reported on a classroom report form that gives individual student and classroom average scores for the six subtests, total math and total reading. In most cases, reports of test results will be mailed to individual schools within 24 hours of receipt of the answer sheets by the scoring agency. If reports are sent to district offices rather than individual schools (at the districts' option), they will be mailed within 48 hours of receipt of the answer sheets from all schools in the district.

Districts may request, prior to the testing, a district averages report and a report of the percent of students answering each test item correctly. District averages will be sent at the same time as statewide averages, probably around May 15th. Reports showing percent of students answering each item correctly will be mailed to each district in September 1985.

7. Interpretation

As in 1981 and 1983, panels of educators from across the state will establish performance expectations for each test item so that apparent strengths and weaknesses in actual results can be identified. This interpretive information will be provided as early in the 1985-86 school year as possible.

If You Have Questions

As always, we would like to hear from you if you have questions or comments about the Alaska Statewide Assessment — either results from the 1983 testing or plans for the 1985 testing. The Department of Education encourages the participation of local educators in the assessment process. By calling the Assessment Hotline, we hope to make that participation as convenient as possible.

Simply call (503)223-3396 **COLLECT** any weekday between 9 a.m. and 4 p.m., Pacific Time. Tell the person answering that you are calling about the Alaska Statewide Assessment. Your call will be accepted and someone knowledgeable about the assessment will speak with you. It's as easy as that.

Assessment Hotline – (503)223-3396

TESTING AND CURRICULUM

A Variety of Perspectives

Almost a year ago, the Department of Education's Offices of School Improvement and Research & Evaluation sponsored the development of position papers written by five professional teachers' associations. The reading and mathematics position papers have particular relevance for persons interested in the results of the 1983 Alaska Statewide Assessment, for there is substantial overlap between the recommendations in those position papers and the objectives tested in the Assessment.

The Department sponsored these position papers, in part, to provide a resource for local districts in reviewing their own curricula. A related activity will provide additional resources. For the past several months, many Department staff have been spending substantial portions of their time reviewing scope and sequence charts and various other curriculum materials from a wide variety of sources, including the Assessment objectives and the professional association recommendations. This review is culminating in the development of elementary and secondary curriculum guides.

After a year-long review during 1984-85 by educators throughout the state, the Department will disseminate the models. While these **particular** curricula will not be mandated, districts will be responsible for having **some** documented curriculum in place. It's our hope that this work will be helpful as they develop, review or revise their curriculum.

We think it is particularly appropriate to cross reference the recommendations made by the reading and math professional associations with results from the 1983 Assessment. This type of review can provide the information needed to plan technical assistance at the state or local level, be it in-service, special projects or onsite consultation. And that, really, is a primary purpose for testing: to identify areas where students are not doing as well as hoped, and to concentrate instructional efforts so performance is eventually improved. If testing contributes to increased learning, then it has served its purpose well.

Darby Anderson
Office of School Improvement

Bob Silverman
Office of Research & Evaluation

Office of Research & Evaluation
Alaska Department of Education
Pouch F
Juneau, AK 99811

ASSESSMENT REPORT—Please Circulate To All Interested Individuals

ASSESSMENT REPORT 10: An Update on the Alaska Statewide Testing Program

Better Assessment Scores ARE Possible. A Test Author Tells How . . .

In March, nearly 15,000 Alaska students will take the Statewide Assessment Tests—fourth and eighth grade reading and mathematics tests which provide the only statewide indication of how well Alaska students are meeting achievement expectations for the basic skills. Most of those students want to do as well as possible on the test. But even when a district's students want to perform well, local educators may believe that their scores are the result **just** of learning over an extended period and can't really be improved in the short term.

This view is incorrect. As one of the authors of the tests which will be used in March, I know of at least seven simple steps that can improve students' scores. None of the steps make students any "smarter," but they do ensure that students get credit for everything they ***do*** know.

1. Make certain students have sufficient time to complete the test. The fourth and eighth grade tests each have a reading and mathematics section; each section has 45 or fewer items. Most students complete each section within 30-45 minutes, but a few students may still be making progress on the test after nearly an hour. It also takes some time to get the materials passed out and the administration directions read. If a school tries to fit the testing into, say, a 47-minute class period, students' scores will not be as good as they deserve to be. (This happened in one of Alaska's districts in 1983. Looking at the answer sheets showed us that numerous students had run out of time and hadn't attempted the last items on the math section; the district's scores reflected this problem.)

2. Neatness counts. As with any machine-scored test, if an item doesn't appear to have a single correct answer, it has to be counted wrong. The Alaska Assessment has less problem with multiple marks than most other tests; every answer sheet is examined by hand before it is scored. But, even so, we can't always tell what answer the student intended. Unfortunately, an answer that can't be identified is no better than no answer at all. A dark (No. 2) pencil to mark the desired answer and a good eraser to get rid of unwanted answers are important tools for all students taking the test.

3. Encourage students to guess. Research has consistently shown that students who "guess" at an answer are rarely just making a random selection. More often, they can eliminate one or two of the options in a multiple choice question, thereby having to guess among a smaller number of options. The probability of getting the item correct is thus much better than chance and reflects that students really do know something about the item. Even students who have just the slightest idea should be encouraged to guess. In the assessment scoring, a blank response is always wrong; a "guess" has a chance of being right.

4. Make certain that students take the test seriously. The smaller a school or district, the greater the effect that any one student has on the group average. It takes a lot of really good performances to make up for the low score of a single student who just marked the answer sheet randomly. Districts hoping to obtain the best scores possible should alert the teachers administering the test to verify that students are conscientiously working through the items.

(continued on next page)

Improving Test Scores Is Sometimes Simple

(Continued from page 1)

5. Do the practice items. About a quarter of the teachers administering the 1983 test didn't go through the practice items with their students. They figured, usually correctly, that their students already knew how to take a multiple choice test. But the practice items also alert the student to an unusual format for the math application items in the test. Students who do not go through the practice items may encounter some difficulty with these items.

6. Know what's on the test. The content of the 1985 Assessment has already been provided to schools in the "Road Map" poster sent at the beginning of the school year. In addition, the 1981 and 1983 tests have items and objectives very similar to the 1985 tests, and copies of those tests are available for your review. We're not suggesting that any school or district change their curriculum or instruction just because an item is on the assessment. But when there is an overlap between the skills tested in the assessment and the district curriculum, the students should know how to answer items like those contained on the assessment.

7. Strengthen the weaknesses. Following the 1983 assessment, the Alaska Council of Teachers of Mathematics and the Alaska State Reading Association prepared position papers giving their expectations for student performance. Based on those expectations, assessment staff identified student weaknesses—areas where the performance of the state's students wasn't up to the goals of the state's teachers. Those weak objectives are noted in the Road Map of test content. Because the skills cited are important skills in the eyes of the professional organizations, they are good areas to address when preparing students for the assessment.

As noted, none of these suggestions does anything to change the underlying achievement level of the students. But they do point out ways to ensure that the performance of district students is as good as it deserves to be.

Michael D. Hiscox
Interwest Applied Research

District Testing Representatives Listed . . .

While staff at the Department of Education work hard to make the administration of the Alaska Statewide Assessment as smooth as possible, it would not be feasible to conduct the testing without the cooperation of Alaska's school and district office staff. Primary among those helping to make testing logistics run smoothly are the District Testing Representatives, district staff who serve as liaisons between the district and the Department and its contractor.

In addition to bringing information about testing needs to the Department's attention, the testing representatives also serve as a resource when teachers and other administrators in the district have questions about the Statewide Assessment.

A listing of the 1984-85 District Testing Representatives appears below. Staff with questions about the assessment should contact the appropriate individual(s) in their district.

- Beverly Raphel, Adak
- Ralph Rudzik, Alaska Gateway
- Dick Anderson, Aleutian Region
- Fred Stofflet, Anchorage
- Walter Bromenschenkel, Annette Island
- Mary Rubadeau, Bering Strait
- Len Mundorf, Bristol Bay
- Duane Fryrear and Ron Gleason, Chatham
- Dennis Hoyt, Chugach
- Mike Johnson, Copper River
- Irene Fairall-Gunnerson, Cordova
- John DeLong, Craig
- Doug Whitener, Delta-Greely
- Wes Hill, Dillingham
- Mary Francis, Fairbanks North Star
- Gail Higgins, Galena
- Allen Heinrich, Haines
- Kenneth Dettering, Hoonah
- Richard Gigo, Hydaburg
- Mike Baumgartner, Iditarod
- David Thomas, Juneau
- Charlotte Conner, Kake
- Dennis Daggett, Kenai Peninsula
- Ron Hamberger, Ketchikan Gateway
- Lorelei Mack, King Cove
- Bill Denkinger, Klawock
- Dennis Nicholson, Kodiak Island
- Bill Wagnon, Kuspuk
- Jean Burns-Kateley, Lake and Peninsula
- Laura Bernhard, Lower Kuskokwim
- Roger Adams, Lower Yukon
- Ray Carter, Mat-Su
- Mark Manuel, Nenana
- Nancy Schave, Nome
- Walter Berard, North Slope
- Michael Carson, Northwest Arctic
- Ralph Allen, Pelican
- John McCabe, Petersburg
- Rick Luthi, Pribilofs
- Kathy Formella, Railbelt
- Herb Oba, St. Mary's
- Cy Crawley, Sand Point
- Terry Coon, Sitka
- David Lannigan, Skagway
- Dave Dossett, Southeast Island
- Christopher Robinson, Southwest Region
- Ron DeLay, Tanana
- Robert Mutch, Unalaska
- Alice Bosshard, Valdez
- Robert Prunella and Dick Barton, Wrangell
- Stan Vogel, Yakutat
- Robert Cook, Yukon Flats
- Niki McCurry, Yukon/Koyukuk

Our thanks to these individuals for their assistance in the assessment testing. Be sure to give your district's representative a call if you have any questions about the Statewide Assessment.

PROFILE –Educational Testing in Alaska

Anchorage Schools Work to Integrate Testing and Curriculum

In this issue of the *Assessment Report* newsletter, we introduce a new feature to highlight educational testing and evaluation projects of particular interest around the state. At the same time, we'll also introduce some of the educators who serve as District Testing Representatives. In this first article, we interview Dr. Fred Stofflet, Special Assistant to the Superintendent in the Anchorage School District. Dr. Stofflet is in his seventh year with the Anchorage Schools. Prior to his current position as Special Assistant to the Superintendent, he was the district's Director of Assessment and Evaluation.

Fred Stofflet grew up in Beaver Dam, Wisconsin, a town about 45 miles east of Madison. He attended both undergraduate and graduate school at the University of Wisconsin in Madison. His graduate major was educational administration, his minor was educational psychology, with specializations in statistics and computer applications. He worked in the Milwaukee (WI) and Norfolk (VA) school districts before moving to Anchorage. When not busy with testing and evaluation activities in the district, he enjoys fishing and working with youth athletic programs. This past summer he coached a very successful girls' softball team, and he's now coaching a girls' basketball team.

For the past year, Fred has been very involved in the development of criterion referenced tests for the district. Several years ago, educators in the Anchorage School District determined that norm referenced tests just were not providing all the information they needed to shape their instructional programs effectively. Through their efforts, the district launched a major criterion based testing program, under Fred's direction and guidance. In this first interview, Fred shares with us some details of how the Anchorage criterion based testing system was developed, and his hopes for its future potential.

REPORTER: How did the Anchorage testing system get started?

STOFFLET: We wanted something that would give us an opportunity to really improve instruction. The district curriculum people had been working for some time on a comprehensive scope and sequence for grades 3, 6 and 8—in math, language arts (including reading and writing), science, and social studies. I don't want to say it was "finalized"—I don't know that anybody ever reaches that point—but about two years ago, it got to the point where we felt it was pretty much . . .

REPORTER: It was looking the way you wanted to see it look?

STOFFLET: Yes, O.K., I would say that—with the understanding that we're continuing to revise and update it. So, at that point, we asked ourselves, "O.K., what would you say a kid would really have to know at each grade level in each of these subjects?" And the answers we came up with became our primary objectives—sort of "exit goals," if you will, for each grade in question.

REPORTER: And those objectives became the basis for your tests?

STOFFLET: Yes. We hired an outside test writer—a specialist—who came into Anchorage to meet with staff, meet with members of our curriculum committees for each subject. We have a standing committee for each subject area that includes not only professional educators but also specialists—noneducators—in each field.

REPORTER: And did those committee people review the test items?

STOFFLET: They conducted a thorough review in which they asked two questions: First, do the test items cover material actually included in the curriculum? And second, are the test items on level—are they right for the students who'll be taking the tests?

REPORTER: What information came out of that review?

STOFFLET: We had some reading level problems that needed correcting, and also some problems with the science items. They seemed too factual in some instances—they didn't get at the scientific process itself, and that was what we were really wanting to test, more than the recall of facts. We did a field test on the second version of the items in May.

REPORTER: How did you feel about the results of that field testing?

STOFFLET: Well, we would have liked to see better performance—that's usual in testing, of course. And particularly with the science items, we were still concerned about the match between the test and what we're trying to teach. Right now we're in an updating phase. This time around, we're using inhouse staff to write and revise items. It's a difficult decision, you know—who's the best qualified to write your test items? When you pull in a specialist from outside, you may get items that are technically better, but are perceived to be less valid than those written by staff who are more familiar with the curriculum.

"It's a difficult decision, you know—who's the best qualified to write your test items?"

REPORTER: So maybe by having your own staff revise the work of your outside consultant you'll end up with the best of both worlds?

STOFFLET: I see that as a real possibility, yes.

(continued on next page)

PROFILE
Fred Stofflet, Testing Representative
The Anchorage Public Schools

REPORTER: What impact do you project this criterion referenced testing could have on the Anchorage curriculum?
STOFFLET: The potential for impact is there because the criterion referenced model is directly related to our curriculum. It can identify for us the specific areas where we're not doing as well as we'd like, and that's the information we need to improve our instruction.
REPORTER: Do you have some idea now of how strong public support might be for the new testing system?
STOFFLET: We conduct a community survey every couple of years, and on the last one we included a statement — something to the effect that the district annually tests students in grades 1-12 and compares performance to that of other students around the country — do you find this information of value? And parents said definitely yes. Then we asked, "What if we conducted a districtwide criterion referenced testing to see how well our students were doing on objectives we ourselves identified as important: How valuable would you find that information?" And again, the parents' overwhelming response was yes, we would find that information valuable. And then — now here's the really interesting question — we said, "Well, what if you had to choose between the two? Which approach would you prefer?" And the majority — by a fair amount — said we'd rather know how students are doing as measured by the local curriculum, the local objectives. Those results tell us that we're definitely on the right track as far as what the public would like to know about how students are doing.
REPORTER: How about teacher support? Is that strong as well?
STOFFLET: It seems to be fairly strong at this point, yes. And I'd say that will increase as we improve our items and get a closer match between the test and what is actually being taught in the classrooms.
REPORTER: If you had to summarize your feelings about this testing approach, what would you say?
STOFFLET: Personally, I see a need for a mixed kind of testing model. Politically and realistically, we need to compare ourselves to the outside world. But at the same time, we have to realize that 99% of our "educational decisions" are based on information that comes out of a criterion referenced system. So what we need is a balance — not for one to replace the other, but for the two kinds of testing to work together — providing us a comprehensive picture of student performance.

If your district or school is developing a new testing program or makes innovative use of educational testing, and you would like to share that information with fellow educators around Alaska, contact Al Hazelton, Director of Assessment for the Department of Education in Juneau at 465-2966. As space permits, we would like to include your information in a future Assessment Report newsletter.

A Reminder — Weaknesses Shown in the Last Assessment

Concurrent with the Spring 1983 Assessment testing, members of the Alaska State Reading Association prepared a position paper on the reading content important for Alaska students to master. The Assessment Program staff compared those recommendations with performance on the 1983 testing and results from the statewide Interpretive Panel, and identified these areas of weakness. Items measuring each of these areas will be on the 1985 Assessment as well.

Grade 4
- Using an index
- Using an encyclopedia
- Using a table of contents
- Knowing where to find information
- Following written directions
- Consonant blends
- Vowel sounds
- Regular and irregular noun plurals
- Use of the dictionary

Grade 8
- Making a judgment based on evidence
- Recognizing bias
- Using an index
- Plural possessives
- Prefixes

The Alaska Council of Teachers of Mathematics prepared a similar position paper describing important math content. The Assessment Program staff used the same procedures to identify this list of student weaknesses, as measured by the 1983 testing. As with reading, similar content will appear on the 1985 tests.

Grade 4
- Solving mathematical word problems
- Determining how to solve a word problem
- Multiplying with regrouping
- Dividing both with and without remainders
- Application of standard measures (for example, feet in a yard)

Grade 8
- Solving mathematical word problems
- Determining how to solve the problem
- Subtracting and multiplying negative numbers
- Dividing fractions
- Numeric sequences, especially involving negative numbers
- Triangle properties
- Characteristics of pi

BEST COPY AVAILABLE

Helpful Assessment Materials Are Available to Schools

The Alaska Department of Education has numerous resources available to help students and staff with the 1985 Alaska Statewide Assessment Test. This sheet lists those resources, and acts as an order form for schools wishing to request a copy of any of the materials. Check the appropriate boxes and we'll see that the materials are sent right out.

Road Map to the Alaska Statewide Assessment — Matches item content (in the form of simplified statements of objectives) with item numbers in the 1981, 1983 and 1985 tests. Also highlights the content identified as areas of weakness by state professional teachers' associations (see description of *Assessment Report 9* below). ☐

Assessment Report 7 — Provides results of the 1983 Assessment: statewide average scores for the subtests, comparisons of scores for different subgroups of students (males and females, Natives and non-Natives, and so on); answers important questions about interpreting the results. ☐

Assessment Report 8 — Lists student strengths and weaknesses on the 1983 Assessment based on a review of the test content by 70 educators from throughout the state who were nominated by their districts. ☐

Assessment Report 9 — Compares results from the 1983 Assessment with curricular emphases recommended by the Alaska Council of Teachers of Mathematics and the Alaska State Reading Association. ☐

1983 Grade 4 Assessment Test — Includes both the reading and mathematics subtests; content very similar to the 1985 fourth grade test. ☐

1983 Grade 8 Assessment Test — Includes both the reading and mathematics subtests; content very similar to the 1985 eighth grade test. ☐

1981 Grade 4 Assessment Test — Includes both the reading and mathematics subtests; slightly different content from the 1985 fourth grade test (as noted on the *Road Map*), but still good practice for the 1985 test. ☐

1981 Grade 8 Assessment Test — Includes both the reading and mathematics subtests; slightly different content from the 1985 eighth grade test (as noted on the *Road Map*), but still good practice for the 1985 test. ☐

One copy of each of the above materials will be sent to schools upon request. Simply place a check mark in the box next to each item you would like to receive. Then complete the box below, tear off this entire sheet, fold as indicated, and mail. Postage has already been paid.

PRINT CAREFULLY — This is your mailing label.

Name _____

School _____

Address/Box Number _____

City/Town/Village _____

State and Zip Code _____

Fold Here and Tape at Bottom to Seal

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS PERMIT NO. A-1296 PORTLAND, OR

POSTAGE WILL BE PAID BY ADDRESSEE

ALASKA ASSESSMENT MATERIALS
INTERWEST APPLIED RESEARCH
P.O. BOX 8189
PORTLAND, OREGON 97207

ASSESSMENT POLICY

Questions and Answers

As long-time readers of this newsletter know, the Alaska Statewide Assessment has operated under policies adopted nearly ten years ago by the Department of Education and the State Board of Education. As the testing system has matured, however, some of the original policies have been called into question. The questions below address some of these policy issues.

Q. I've heard that there isn't any way to tell how my district did on the assessment compared to other districts in the state. Is this true?

A. It's partly true. A district will know, for example, that its math computation scores were in the top quarter of all the state's districts. But they wouldn't know how the district next door did. When the assessment started, the State Board approved a policy prohibiting such comparisons. This policy has been questioned by several people—including some state legislators and current Board members—so it may be changed at some point. But for now, the original policy stands.

Q. But aren't there statistical reasons why comparisons shouldn't be made?

A. In several cases, Alaska schools and districts have too few students to yield reliable results. Comparing such scores could be misleading. There are two schools of thought. Some say that no information should be made available unless it is completely accurate. Others say that information that's probably correct is better than no information at all. Neither group is wrong; the proper approach depends on the care with which interpretations of the scores are made.

Q. The assessment currently tests fourth and eighth graders every two years in reading and math. What are the chances of getting some more grades and subjects tested, or testing every year? That would help the assessment play a bigger role in our district testing program.

A. Whether to expand the assessment depends largely on how much use will be made of the data. When the assessment started, general statewide information was the main goal. But the recent efforts in effective schooling and curriculum alignment have increased the utility of assessment data at both the state and local level. Whether this potential is enough to warrant expanding the program isn't clear.

Q. You've got some important questions here. How do you intend to find the answers?

A. Beginning in February, an extensive evaluation will be conducted to determine future policies for the assessment. State Board members, legislators, and local and Department of Education staff will be asked to decide what information they need from an assessment and how that information would improve their decision making ability. And every district, through its Testing Representative, will have a chance to comment on proposed changes. By the end of June, a plan for the Assessment Program—perhaps changed from it current form, perhaps not—will be ready for approval.

Q. I've already got some suggestions for improving the assessment. Is there some way that I can get my ideas to you now?

A. Certainly. You can use the Assessment Hotline (503/223-3396) or call Al Hazelton, the Department of Education's Director of Assessment, directly at 465-2966. At either number, there will be someone to record your comments and see that they are included in the planning process.

Assessment on Schedule for March 25-29 Administration

Although the first Alaska Statewide Test was given almost nine years ago, it still takes careful planning to ensure that the tests are of high quality and their administration goes smoothly. In fact, work began nearly a year ago on preparations for the March 1985 administration, with the drafting of new test items designed to match the content tested in the 1981 and 1983 assessments. The items were reviewed by members of the Reading and Mathematics Content Panels—curriculum specialists from local districts and Alaska universities—who offered numerous suggestions for improvements.

The items were revised over the summer and then pilot tested in October. Analysis of the pilot test revealed a few items that needed further refinement, and those revisions were made before the tests went to the printer in early January.

The enrollment information necessary to allow the Department's contractor to prepare testing materials in the appropriate packages for each classroom in the state is also being collected. When the testing materials arrive in the schools in late February or early March, they will be subdivided into classroom packages based on the enrollment information received from the District Testing Representatives.

In some cases, the Testing Representative has requested that testing materials for all schools in the district be sent to the district office for distribution; in other cases, materials will be sent directly to the individual school buildings. Similarly, when classroom reports are prepared, most will be sent directly to the schools so there is minimal turnaround time. In a few cases, though, Testing Representatives will receive the classroom reports and then distribute them to the schools in the district.

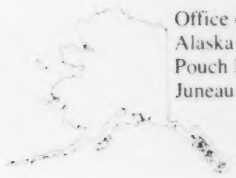
Local staff who have special materials needs or who have any conflicts with the March 25-29 testing period should contact their District Testing Representative (see page 2) or the staff on the Assessment Hotline (503/223-3396 Collect) immediately so that special arrangements can be made.

Questions About the Assessment?

Don't forget that there are staff available to answer any questions you might have about the Alaska Statewide Assessment. One resource to contact is the District Testing Representative for your district, whose name is listed on page 2 of this newsletter. If that person can't answer your question, he or she will contact either Al Hazelton at the Department of Education or staff at Interwest Applied Research, the Department's contractor for the Statewide Assessment.

If you have questions, comments or suggestions about the assessment, you should also feel free to call Interwest directly on the Assessment Hotline. Call collect anytime between 8 a.m. and 3 p.m. Alaska time and tell the operator that you are calling from Alaska about the Statewide Assessment.

Assessment Hotline
(503) 223-3396, Collect



Office of Education, Assessment & Research
Alaska Department of Education
Pouch F
Juneau, AK 99811

ASSESSMENT REPORT—Please Circulate To All Interested Individuals

BEST COPY AVAILABLE